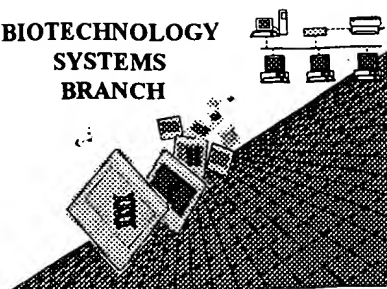


RAW SEQUENCE LISTING **ERROR REPORT**

BIOTECHNOLOGY
SYSTEMS
BRANCH



TECH CENTER 1600/2900

JAN 07 2002

RECEIVED

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/657,986

Source: 1600

Date Processed by STIC: 12/13/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:

<http://www.uspto.gov/web/offices/pac/checker>

ERROR DETECTED**SUGGESTED CORRECTION**SERIAL NUMBER: 09/657,986

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics
 Wrapped Aminos The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 Invalid Line Length The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 Misaligned Amino
 Numbering The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 Non-ASCII The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 Variable Length Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 PatentIn 2.0
 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 Skipped Sequences
 (OLD RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) ...
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 Skipped Sequences
 (NEW RULES) Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000
- 9 Use of n's or Xaa's
 (NEW RULES) Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10 Invalid <213>
 Response Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or Artificial Sequence
- 11 Use of <220> Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 PatentIn 2.0
 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 Misuse of n n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

1600

RAW SEQUENCE LISTING

DATE: 12/13/2001

PATENT APPLICATION: US/09/657,986

TIME: 09:54:28

Input Set : A:\Sequence_listing.asc

Output Set: N:\CRF3\12132001\I657986.raw

delete at beginning of file

W--> 1 4 <110> APPLICANT: Madison, Edwin L.
 5 Semple, Joseph Edward
 6 Coombs, Gary Samuel
 7 Reiner, John Eugene
 8 Ong, Edgar O.
 9 Araldi, Gian Luca
 11 <120> TITLE OF INVENTION: Inhibitors of Serine Protease Activity of Matriptase or
 12 MTSP1
 14 <130> FILE REFERENCE: Corvas 255/049
 16 <140> CURRENT APPLICATION NUMBER: 09/657,986
 17 <141> CURRENT FILING DATE: 2000-09-08
 19 <160> NUMBER OF SEQ ID NOS: 10
 21 <170> SOFTWARE: PatentIn version 3.0

Does Not Comply
Corrected Diskette Needed

ERRORED SEQUENCES

23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 1452
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Homo_sapiens
 28 <400> SEQUENCE: 1

E--> 29 gttgttgggg gcacggatgc ggatgagggc gagtggccct ggcaggtaag
 30 cctgcatgct 60
 E--> 32 caacaacccc cgtgcctacg cctactcccg ctcaccggga ccgtccattc
 33 ggacgtacga 120
 E--> 35 ctggggccagg gccacatctg cgggtgcttc ctcactcttc ccaactggct
 36 ggtctctgcc 180
 E--> 38 gaccgggtcc cgggtgtagac gccacgaagg gagtagagag ggttgaccga
 39 ccagagacgg 240
 E--> 41 gcacactgct acatcgatga cagaggattc aggtactcag accccacgca
 42 gtggacggcc 300
 E--> 44 cgtgtgacga tgtagctact gtctcctaag tccatgagtc tgggggtgcgt
 45 cacctgccgg 360
 E--> 47 ttcttgggct tgcacgacca gagccagcgc agcgcccctg ggggtgcagga
 48 gcgcaggctc 420
 E--> 50 aaggaccgga acgtgctggt ctcggtcgcg tcgcggggac cccacgtcct
 51 cgcgtccgag 480
 E--> 53 aagcgcacga tctcccaccc cttcttcaat gacttcacct tcgactatga
 54 catcgcgctg 540
 E--> 56 ttgcgctagt agagggtggg gaagaagtta ctgaagtgga agctgatact
 57 gtagecgcgac 600
 E--> 59 ctggagctgg agaaaccggc agagtacagc tccatggtgc ggcccatctg
 60 cctgccggac 660
 E--> 62 gacctcgacc tctttggccg tctcatgtcg aggtaccacg ccgggtagac
 63 ggacggcctg 720

(see item 1
on Error Summary
sheet)

global
format error

60

RAW SEQUENCE LISTING

DATE: 12/13/2001

PATENT APPLICATION: US/09/657,986

TIME: 09:54:28

Input Set : A:\Sequence_listing.asc

Output Set: N:\CRF3\12132001\I657986.raw

```

E--> 65 gcctcccatg tcttcctgc cggcaaggcc atctgggtca cgggctgggg
66 acacacccag 780
E--> 68 cggaggggtac agaagggacg gccgttccgg tagaccaggt gcccgacccc
69 tgtgtgggtc 840
E--> 71 tatggaggca ctggcgcgct gatcctgcaa aagggtgaga tccgcgtcat
72 caaccagacc 900
E--> 74 atacctccgt gaccgcgcga ctaggacgtt tccccactct aggcgcagta
75 gttggtctgg 960
E--> 77 acctgcgaga acctcctgcc gcagcagatc acgccgcgca tgatgtgctg
78 gggcttcttc 1020
E--> 80 tggacgtctt tggaggacgg cgtcgtctag tgcggcgcgt actacacgca
81 cccgaaggag 1080
E--> 83 agcggcggcg tggactcctg ccagggtgat tccgggggac ccctgtccag
84 cgtggaggcg 1140
E--> 86 tcgccgcgcg acctgaggac ggtcccacta agggccctcg gggacaggtc
87 gcacctccgc 1200
E--> 89 gatgggcgga tcttcaggc cgggtgtggtg agctggggag acggctgcgc
90 tcagaggaac 1260
E--> 92 ctaccgcctt agaaggctcg gccacaccac tcgacccctc tgccgacgcg
93 agtctccttg 1320
E--> 95 aagccaggcg tgtacacaag gctccctctg ttccgggact ggatcaaaga
96 gaacactggg 1380
E--> 98 ttcggtccgc acatgtgttc cgaggagagac aaagccctga cctagtttct
99 cttgtgaccc 1440
E--> 101 gtatagcata tc
102 1452
196 <210> SEQ ID NO: 5
197 <211> LENGTH: 28
198 <212> TYPE: DNA
199 <213> ORGANISM: Homo_sapiens
201 <400> SEQUENCE: 5
E--> 202 cacccttctt tcaatgactt caccttcg
203 28
206 <210> SEQ ID NO: 6
207 <211> LENGTH: 18
208 <212> TYPE: DNA
209 <213> ORGANISM: Homo_sapiens
211 <400> SEQUENCE: 6
E--> 212 tacctctcct acgactcc
213 18
216 <210> SEQ ID NO: 7
217 <211> LENGTH: 25
218 <212> TYPE: DNA
219 <213> ORGANISM: Homo_sapiens
221 <400> SEQUENCE: 7
E--> 222 gaggttctcg cagggtgtct ggttg
223 25
226 <210> SEQ ID NO: 8
227 <211> LENGTH: 39

```

*same error**same**same**same*

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/657,986

DATE: 12/13/2001

TIME: 09:54:28

Input Set : A:\Sequence_listing.asc

Output Set: N:\CRF3\12132001\I657986.raw

228 <212> TYPE: DNA
229 <213> ORGANISM: Homo_sapiens
231 <400> SEQUENCE: 8
E--> 232 ctcgagaaaa gagttgttgg gggcacggat gcggatgag
233 39
236 <210> SEQ ID NO: 9
237 <211> LENGTH: 36
238 <212> TYPE: DNA
239 <213> ORGANISM: Homo_sapiens
241 <400> SEQUENCE: 9
E--> 242 gcggccgcac tataccccag tgttctcttt gatcca
243 36

*same**same*

see next 2 pages for more error

09/657,986

4

His Asp Gln Ser Gln Arg Ser Ala Pro Gly Val Gln Glu Arg Arg Leu
 65 70 75 80
 Lys Arg Ile Ile Ser His Pro Phe Phe Asn Asp Phe Thr Phe Asp Tyr
 85 90 95
 Asp Ile Ala Leu Leu Glu Leu Glu Lys Pro Ala Glu Tyr Ser Ser Met
 100 105 110
 Val Arg Pro Ile Cys Leu Pro Asp Ala Ser His Val Phe Pro Ala Gly
 115 120 125
 Lys Ala Ile Trp Val Thr Gly Trp Gly His Thr Gln Tyr Gly Gly Thr
 130 135 140
 Gly Ala Leu Ile Leu Gln Lys Gly Glu Ile Arg Val Ile Asn Gln Thr
 145 150 155 160
 Thr Cys Glu Asn Leu Leu Pro Gln Gln Ile Thr Pro Arg Met Met Cys
 165 170 175
 Val Gly Phe Leu Ser Gly Gly Val Asp Ser Cys Gln Gly Asp Ser Gly
 180 185 190
 Gly Pro Leu Ser Ser Val Glu Ala Asp Gly Arg Ile Phe Gln Ala Gly
 195 200 205
 Val Val Ser Trp Gly Asp Gly Cys Ala Gln Arg Asn Lys Pro Gly Val
 210 215 220
 Tyr Thr Arg Leu Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr Gly
 225 230 235 240
 Val

<210> 3
 <211> 23
 <212> PRT
 <213> Homo_sapiens

<220>
 <221> misc feature
 <223> R=A,G; V=G,A,C; W=A,T; S=G,C; Y=C,T; H=A,T,C

<400> 3

Thr Gly Gly Arg Thr Ile Val Thr Ile Trp Ser Ile Gly Cys Ile Arg
 1 5 10 15
 Cys Ile Cys Ala Tyr Thr Gly
 20

<210> 4
 <211> 30
 <212> PRT
 <213> Homo_sapiens

<220>

These only
apply to nucleotide
sequences.

Sequence 3
is an
amino acid
sequence.

<221> misc feature

<223> R=A,G; V=G,A,C; W=A,T; S=G,C; Y=C,T; H=A,T,C

same error - these
only
apply
to
nucleotide
sequences

<400> 4

Ile Gly Gly Ile Cys Cys Ile Cys Cys Ile Ser Trp Arg Thr Cys Ile
1 5 10 15

Cys Cys Tyr Thr Ile Arg Cys Ala Ile Gly His Arg Thr Cys
20 25 30

<210> 5

<211> 28

<212> DNA

<213> Homo_sapiens

<400> 5

caccccttct tcaatgactt caccttcg
28

<210> 6

<211> 18

<212> DNA

<213> Homo_sapiens

<400> 6

tacctctcct acgactcc
18

<210> 7

<211> 25

<212> DNA

<213> Homo_sapiens

<400> 7

gaggttctcg caggtggtct gggttg
25

<210> 8

<211> 39

<212> DNA

<213> Homo_sapiens

<400> 8

ctcgagaaaa gagttggttg gggcacggat gcggatgag
39

<210> 9

<211> 36

<212> DNA

<213> Homo_sapiens

<400> 9

gcggccgcac.tataccccag tggttctcttt gatcca
36

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/657,986

DATE: 12/13/2001

TIME: 09:54:29

Input Set : A:\Sequence_listing.asc

Output Set: N:\CRF3\12132001\I657986.raw

L:1 M:259 W: Allowed number of lines exceeded, (1) GENERAL INFORMATION:

L:29 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:50 SEQ:1

M:254 Repeated in SeqNo=1

L:202 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:28 SEQ:5

L:212 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:18 SEQ:6

L:222 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:25 SEQ:7

L:232 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:39 SEQ:8

L:242 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:36 SEQ:9